A Conformity Test for Cointegration

by
Phoebus J. Dhrymes, Columbia University

April
~~March~~ 1996

Discussion Paper Series No. 9596-10

pages: 32

A Conformity Test for Cointegration

PHOEBUS J. DHRYMES *
Columbia University

October 1995; revised April 1996

Abstract

This paper formulates a conformity test for cointegration for a multivariate $I(1)$ process obeying a VAR specification. The test statistic is a function of the characterisitic roots of the sample covariance matrix of the cointegral vector; the latter is obtained from the unrestricted estimator of the underlying parameters of the VAR. It is further shown that this test procedure is applicable to the case where the $I(1)$ process is a $MIMA(k)$, i.e. a multivariate integrated moving average process, the moving average being of order $k < \infty$.

The test statistic, under the null of cointegration, has a normal limiting distribution.

Key Words: Cointegration, Cointegration test, characteristic roots, VAR, $MIMA(k)$.

1 Introduction and Summary

Let $\{X_{t\cdot} : t \in \mathcal{N}\}$ be a stochastic sequence defined on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$. If the sequence is taken to be $I(1)$, in the sense that $\{(I-L)X_{t\cdot} : t \in \mathcal{N}\}$ is **strictly stationary**, the question often arises as to whether the sequence in question is cointegrated. The latter means, in this context, that there exists a $q \times r$ matrix B of rank $r \leq q$ such that $X_{t\cdot}B$ is (strictly) stationary. A number of tests have been proposed in the literature, some formal, some informal. The somewhat informal tests involve running a regression of one of the elements on the others and using the Dickey-Fuller test, Dickey and Fuller (1979), (1981), to test the hypothesis of cointegration; other informal tests are also given in Engle and Granger (1987), Stock and Watson (1988), *inter alios.*

* I wish to express my appreciation to Dimitrios Thomakos for his very able research assistance.

More formal tests are given in Phillips and Ouliaris (1990), and Johansen (1988), (1991), to mention but a few. All such tests employ an indirect approach in that they explore an implication of the cointegration hypothesis beyond the property that constitutes its definition.

In this paper we explore a conformity test for cointegration, and give the limiting distribution of the test statistic. We place the discussion in the VAR context of cointegration popularized by Johansen as noted above, but the results are equally applicable to contexts that are less constrained. Finally, the test statistic is shown to be asymptotically normal; thus, tests for cointegration may be carried out in standard fashion, in contrast with other procedures that require special tabulations.

2 Notation and Problem Formulation

Consider the standard VAR

$$X_{t\cdot}\Pi(L) = \sum_{j=0}^{n} X_{t-j\cdot}\Pi_j = \epsilon_{t\cdot}, \quad t \geq 1, \quad \Pi_0 = I_q, \quad X_{-t\cdot} = 0, \quad t \geq 0, \tag{1}$$

where $X_{t\cdot}$ is a q-element **row** vector, the error process being a $MWN(\Sigma)$, i.e. a multivariate white noise[1] process with mean zero and covariance matrix $\Sigma > 0$; normality is not necessary, as in the case of Johansen (1988), (1991).

"Dividing" $\Pi(L)$ by $(I - L)$, where L is the usual lag operator we find, after some rearrangement,

$$(I - L)X_{t\cdot} = -X_{t-1\cdot}\Pi(1) + x_{t\cdot}\Pi^* + \epsilon_{t\cdot}, \tag{2}$$

where

$$\begin{aligned} x_{t\cdot} &= (\Delta X_{t-1\cdot}, \Delta X_{t-2\cdot}, \ldots, \Delta X_{t-n+1\cdot}), \quad t = 1, 2, \ldots T, \\ \Pi^* &= (\Pi_1^{*\prime}, \ldots, \Pi_{n-1}^{*\prime})', \quad \Pi_j^* = \sum_{i=j+1}^{n} \Pi_i, \quad \Pi(1) = \sum_{j=0}^{n} \Pi_j. \end{aligned}$$

If the process is not cointegrated, $\Pi(1)$ is the zero matrix; if the process is actually stationary, $\Pi(1)$ is nonsingular; if $\Pi(1)$ is **singular**, but **nonnull**, the process is cointegrated. The rank r of $\Pi(1)$ is said to be the cointegration rank.

By the rank factorization theorem, see Dhrymes (1984) p. 23, there exist matrices Γ, B both of dimension $q \times r$ and rank r such that

$$\Pi(1) = B\Gamma'. \tag{3}$$

[1] Normally, the term "white noise" denotes a sequence of zero mean orthogonal vectors, but in this paper we shall take it to denote a sequence of zero mean i.i.d. random vectors.

The intuition and general conceptual framework of the conformity test for cointegration is as follows. In Eq. (2) we estimate by least squares $\Pi(1)$ without the restriction in Eq. (3), say $\hat{\Pi}(1)$, and form the sample covariance matrix

$$\hat{M} = \frac{1}{T}\hat{\Pi}'(1)P_{-1}'P_{-1}\hat{\Pi}(1), \quad P_{-1} = (X_{t-1\cdot}), \quad t = 1, 2, \dots T. \tag{4}$$

This approach exploits the fact that, under the stated conditions,

$$T^{\alpha}\hat{\Pi}(1) \xrightarrow{\mathrm{P}} -\Pi(1)$$

for any $\alpha \in [0, 1)$. Thus,

$$\hat{M} \sim (1/T)\Pi(1)'P_{-1}'P_{-1}\Pi(1)$$

which, under the null of cointegration, is a transformation of the "sample" covariance matrix of a **strictly stationary process**. More precisely, it will be shown below that

$$\operatorname*{plim}_{T\to\infty} \hat{M} = M_{zz} = \Gamma M_* \Gamma', \quad \operatorname*{plim}_{T\to\infty} \frac{B'P_{-1}'P_{-1}B}{T} = M_*, \tag{5}$$

where M_{zz} is at least a positive semidefinite matrix. Thus, if M_{zz} is **nonsingular**, the X-sequence is $I(0)$. If it is singular, but not the zero matrix, the sequence is cointegrated. Finally, if the hypothesis that M_{zz} is not the zero matrix is rejected, the sequence is $I(1)$, but **not cointegrated**.

To implement this procedure obtain the least squares estimator

$$\begin{aligned}\hat{\Pi}(1) &= -\Pi(1) + (V'V)^{-1}V'U, \quad V = NP_{-1}, \\ N &= I_q - X(X'X)^{-1}X', \quad U = (\epsilon_{t\cdot}).\end{aligned} \tag{6}$$

Define,

$$\begin{aligned}\hat{M} &= \hat{\Pi}(1)'\left(\frac{P_{-1}'P_{-1}}{T}\right)\hat{\Pi}(1) = \left(\frac{\Pi(1)'P_{-1}'P_{-1}\Pi(1)}{T}\right) \\ &\quad + A_{12} + A_{12}' + A_{22} \\ A_{12} &= -\frac{1}{T}\left(\frac{U'NP_{-1}}{T}\right)\left(\frac{P_{-1}'NP_{-1}}{T^2}\right)^{-1}\left(\frac{P_{-1}'P_{-1}\Pi(1)}{T}\right), \\ A_{22} &= \frac{1}{T}\left(\frac{U'NP_{-1}}{T}\right)\left(\frac{P_{-1}'NP_{-1}}{T^2}\right)^{-1}\left(\frac{P_{-1}'P_{-1}}{T^2}\right) \\ &\quad \times \left(\frac{P_{-1}'NP_{-1}}{T^2}\right)^{-1}\left(\frac{P_{-1}'NU}{T}\right),\end{aligned} \tag{7}$$

and note that

$$A_{11} = \left(\frac{\Pi(1)' P_{-1}' P_{-1} \Pi(1)}{T}\right) \xrightarrow{\text{a.c.}} M_{zz}, \quad A_{12} \xrightarrow{\text{d}} 0, \quad A_{22} \xrightarrow{\text{d}} 0, \quad \text{and}$$

$$\lambda I_q - \hat{M} \xrightarrow{\text{d}} \lambda I_q - M_{zz}. \tag{8}$$

Hence, the ordered characteristic roots of $\hat{M}$ converge to the (ordered) characteristic roots of M_{zz}, and it is evident that, as expected, the number of zero roots of $\Pi(1)$, and thus of M_{zz}, corresponds to the number of unit roots of $|\Pi(z)| = 0$, and its rank corresponds to the cointegration rank.

To examine the limiting distribution of such roots, we first note that A_{12}, A_{22} both converge to zero at the rate of T^{α}, for $\alpha \in [0,1)$. Thus, the limiting distribution of the roots depends entirely on the first term, and we obtain

$$\sqrt{T}\left(\hat{M} - M_{zz}\right) \sim \sqrt{T}\left[\left(\frac{1}{T}\Pi(1)' P_{-1}' P_{-1} \Pi(1) - M_{zz}\right)\right], \tag{9}$$

owing to the fact that

$$\sqrt{T} A_{12} \xrightarrow{\text{d}} 0, \quad \sqrt{T} A_{22} \xrightarrow{\text{d}} 0. \tag{10}$$

We note that

$$\frac{1}{\sqrt{T}}\Pi(1)' P_{-1}' P_{-1} \Pi(1) - \sqrt{T} M_{zz} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(z_{t-1\cdot}' z_{t-1\cdot} - M_{zz}\right), \tag{11}$$

where $z_{t-1\cdot} = X_{t-1\cdot}\Pi(1)$, and

$$\left\{z_{t\cdot}' z_{t\cdot} - M_{zz} : t \in \mathcal{N}_+\right\}$$

is a zero mean strictly stationary process. If the $MWN(\Sigma)$ that defines the $VAR(n)$ model is assumed in this context to be normal, the entities of Eq. (11) obey

$$E \parallel z_{t\cdot}' z_{t\cdot} - M_{zz} \parallel^{2+\alpha} < \infty, \quad \alpha > 0. \tag{12}$$

If normality is not assumed the moment condition above needs to be assumed explicitly. It follows, from Corollary 1a in Chapter 5, Dhrymes (1995) that

$$\psi_T = \text{vec}\left(\frac{1}{\sqrt{T}}\Pi(1)' P_{-1}' P_{-1} \Pi(1) - \sqrt{T} M_{zz}\right)$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \text{vec}\left(z_{t-1\cdot}' z_{t-1\cdot} - M_{zz}\right)$$

obeys a CLT. Consequently, its limiting distribution is given by

$$\psi_T \xrightarrow{\mathrm{d}} \psi \sim N(0, \Psi^*), \quad \text{where} \quad \Psi^* = \lim_{T\to\infty} E\psi_T\psi_T'. \tag{13}$$

Evidently, the matrix Ψ^* is **singular** since the vector $\zeta_{t\cdot}$ contains redundancies such as, e.g. $z_{t-1,i}z_{t'-1,j} - m_{ij}(0)$ and $z_{t-1,j}z_{t'-1,i} - m_{ji}(0)$; still, it is preferable to proceed in this formal fashion rather than consider only the distinct elements of ψ_T. We now derive the form of the matrix Ψ^*. By definition

$$\begin{aligned} \Psi_T^* &= E\psi_T\psi_T' = \sum_{t=1}^{T}\sum_{t'=1}^{T} E\zeta_{t\cdot}'\zeta_{t'\cdot}, \\ \zeta_{t\cdot}' &= z_{t-1\cdot}' \otimes z_{t-1\cdot}' - \mathrm{vec}(M_{zz}). \end{aligned} \tag{14}$$

We modify the notation above so that it becomes more flexible. Thus, put

$$M(\tau) = Ez_{t+\tau\cdot}'z_{t\cdot}, \text{ so that } \quad M(-\tau) = M(\tau)'. \tag{15}$$

In this notation what we had called M_{zz} above becomes $M(0)$, and we further define

$$\mathrm{vec}[M(\tau)] = m(\tau), \quad \tau = 0, \pm1, \pm2, \ldots, \tag{16}$$

so that

$$E\zeta_{t\cdot}'\zeta_{t'\cdot} = E\left(z_{t-1\cdot}'z_{t'-1\cdot} \otimes z_{t-1\cdot}'z_{t'-1}\right) - m(0)m(0)'. \tag{17}$$

If the underlying $MWN(\Sigma)$ is normal, or if the fourth order cumulants of this distribution are appropriately behaved, see e.g. Hannan (1970), pp. 208-229, or Anderson (1971), pp. 250-271, we find that the (r,s) element of the (i,j) block of Ψ_T^* is given by

$$\Psi_{T,(i,j),(r,s)}^* = \frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}\left[m_{ij}(t-t')m_{rs}(t-t') + m_{is}(t-t')m_{jr}(t'-t)\right]. \tag{18}$$

Moreover, since $m_{jr}(-\tau) = m_{rj}(\tau)$, we may write

$$\Psi_{T,(i,j),(r,s)}^* = \sum_{\tau=-T+1}^{T-1}\left(1 - \frac{|\tau|}{T}\right)\left[m_{ij}(\tau)m_{rs}(\tau) + m_{is}(\tau)m_{rj}(\tau)\right]. \tag{19}$$

If the series above converges, as it will under the standard assumptions of this literature, we conclude

$$\begin{aligned} \Psi_T^* &= \sum_{\tau=-T+1}^{T-1}\left(1 - \frac{|\tau|}{T}\right)[M(\tau)\otimes M(\tau) + m_{\cdot j}(\tau)m_{\cdot i}(\tau)'] \\ &\to \Psi^* = \sum_{\tau=-\infty}^{\infty}[M(\tau)\otimes M(\tau) + m_{\cdot j}(\tau)m_{\cdot i}(\tau)']. \end{aligned} \tag{20}$$

Noting that

$$[m_{\cdot j}(\tau)m_{\cdot i}(\tau)'] = \begin{bmatrix} M(\tau)\otimes m_{\cdot 1}(\tau)' \\ M(\tau)\otimes m_{\cdot 2}(\tau)' \\ \vdots \\ M(\tau)\otimes m_{\cdot q}(\tau)' \end{bmatrix}, \tag{21}$$

we may verify that each term

$$M(\tau)\otimes M(\tau) + \begin{bmatrix} M(\tau)\otimes m_{\cdot 1}(\tau)' \\ M(\tau)\otimes m_{\cdot 2}(\tau)' \\ \vdots \\ M(\tau)\otimes m_{\cdot q}(\tau)' \end{bmatrix}$$

contains certain row redundancies, as was to be expected. Since we are dealing with real processes, $M(\tau) = M(-\tau)'$. Consequently, we may rewrite

$$\Psi^* = M(0)\otimes M(0) + [m_{\cdot j}(0)m_{\cdot i}(0)' + m_{\cdot i}(0)m_{\cdot j}(0)'] \tag{22}$$

$$+\sum_{\tau=1}^{\infty}\left[(M(\tau)\otimes M(\tau)) + (M(\tau)\otimes M(\tau))'\right] \tag{23}$$

$$+\sum_{\tau=1}^{\infty}\{[m_{\cdot j}(\tau)m_{\cdot i}(\tau)'] + [m_{\cdot i}(\tau)m_{\cdot j}(\tau)']\}.$$

A simpler representation of the covariance matrix is obtained by noting that since the ζ-sequence is strictly stationary and square integrable

$$\Psi^* = 2\pi f(0), \tag{24}$$

where, provided it is continuous at zero, $f(\lambda)$ is the spectral matrix of the process [2] $\{\zeta_{t\cdot} : t \geq 1\}$, as the latter is defined in Eq. (14).

We have therefore proved

Theorem 1. In the context of the discussion above, consider the **unrestricted** estimator $\hat{\Pi}(1)$ and the matrix

$$\hat{M} = \frac{1}{T}\hat{\Pi}(1)'P_{-1}'P_{-1}\hat{\Pi}(1).$$

The following statements are true:

[2] Strictly speaking this is required of the spectral matrix of the ζ-process, after redundancies have been eliminated.

i. $\hat{M} \xrightarrow{\mathrm{d}} M_{zz} = M(0)$, [3] and thus in probability as well, by Proposition 40, p. 263 in Dhrymes (1989);

ii. $\mathrm{vec}\{\sqrt{T}[\hat{M} - M(0)]\} \xrightarrow{\mathrm{d}} N(0, \Psi^*)$, where Ψ^* is as defined in Eq. (20), and Eq. (24), f being the spectral matrix of the sequence $\{\zeta_{t\cdot} : t \geq 1\}$, as the latter is defined in Eq. (14).

Remark 1. If the X-sequence is **stationary** all roots of M are positive (and assumed distinct); thus, the covariance matrix Ψ^* of the limiting distribution above, after removal of redundancies, is positive definite. If, on the other hand, some roots of $M(0)$ are zero, as would be the case under cointegration, we note a certain complication. Let Λ and Q be, respectively, the matrices of characteristic roots and characteristic vectors of $M(0)$. If r_0 of the characteristic roots are null, we have the following representation

$$Q'M(0)Q = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \tag{25}$$

where $\Lambda_2 = 0$, i.e. it contains the r_0 zero roots, and $\Lambda_1 > 0$. Partitioning conformably, $Q = (Q_1, Q_2)$ we have that

$$Q_2'M(0)Q_2 = 0; \quad \text{moreover, since} \quad \| M(\tau) \| \leq \| M(0) \|, \quad Q_2'M(\tau)Q_2 = 0, \tag{26}$$

for all τ. Thus, the suitibably modified matrix Ψ^* is **singular**.

We have therefore proved

Corollary 1. Under the null of cointegration, the limiting distribution of $\sqrt{T}\mathrm{vec}[\hat{M} - M(0)]$ has a **singular covariance matrix**.

We are now in a position to formulate a (conformity) test for the presence of cointegration based on the result above. Thus, we have

Theorem 2. Under the conditions of Theorem 1, let γ be a small number, say $\gamma = .001$, then the conformity test for cointegration (CCT) may be formulated as

H_0: $\mathrm{tr}\, M(0) = \gamma$,

as against the alternative,

H_1 : $\mathrm{tr} M(0) > \gamma$.

[3] For simplicity of notation, in the remainder of the paper, we shall use $M(0)$ to denote M_{zz}.

Since

$$\sqrt{T}[\mathrm{tr}\hat{M} - \mathrm{tr}M(0)] \xrightarrow{\mathrm{d}} N(0, \phi^2), \tag{27}$$

the conformity test statistic may be formulated as

$$CCTS : \frac{\sqrt{T}(\mathrm{tr}\hat{M} - \gamma)}{\phi} \xrightarrow{\mathrm{d}} N(0, 1). \tag{28}$$

If the hypothesis H_0 above is accepted, we should conclude that cointegration does not exist and the X-process is $I(1)$ and **not cointegrated**. If it is rejected, we conclude that **it is cointegrated**.

Remark 2. One might ask why we inserted the parameter γ in this discussion. From a practical point of view, it makes no particular difference since in the presence of cointegration $\mathrm{tr}\,\hat{M}$ overwhelms γ. The reason for its insertion, however, is simply to uphold the applicability of the limiting distribution to the null. Notice that if we state the null as $\mathrm{tr}\,M(0) = 0$, it would mean that this **is not an admissible hypothesis**, given the conditions employed in deriving the limiting distribution. On the other hand $\mathrm{tr}\,M(0) = \gamma$ **is, strictly speaking, admissible**. Consequently, this device preserves the formal niceties and at the same time gives us a particularly simple procedure in testing for the presence of cointegration.

We next turn to the question of the rank of cointegration. This necessitates the extraction of the characteristic roots of $\hat{M}$ and, through them, tests of certain hypotheses.

Thus, we shall be dealing with tests for cointegration and/or stationarity. To rule out stationarity we test the null

$H_0 :\ \lambda_{\min} = 0$

as against the alternative

$H_1 :\ \lambda_{\min} > 0$.

Acceptance indicates that the sequence is $I(1)$, and cointegrated of rank at most $q-1$. More generally, if the cointegration rank is r, the $q-r$ smallest characterstic roots of $M(0)$ must be zero. Therefore, to devise such a test we need to obtain the limiting distribution of the characteristic roots of $\hat{M}$. Such a result is not available in the literature. To this end, we have

Theorem 3. In the context of Theorem 1, the (ordered) characteristic roots of $\hat{M}$, contained in the diagonal matrix $\tilde{\Lambda}$ obey

$$\sqrt{T}(\tilde{\Lambda} - \Lambda) \sim d^*[\sqrt{T}Q'[\hat{M} - M(0)]Q],$$

where the notation $d^*[A]$ indicates the **diagonal elements of the matrix** A, and Q is the orthogonal matrix of the decomposition $M(0) = Q\Lambda Q'$, as given in Eq. (25).

Proof: Since $\hat{M}$ and $M(0)$ are at least positive semidefinite, by Proposition 52 in Dhrymes (1984) pp. 61-62, they have the (orthogonal) decomposition

$$\hat{M} = \tilde{Q}\tilde{\Lambda}\tilde{Q}', \quad M(0) = Q\Lambda Q'. \tag{29}$$

Moreover, by the results of Theorem 1, Proposition 28, Corollary 5, pp. 242-244 in Dhrymes (1989), $\tilde{Q}$, $\tilde{\Lambda}$ converge, respectively, to Q and Λ; in addition, $\sqrt{T}(\tilde{Q} - Q)$, and $\sqrt{T}(\tilde{\Lambda} - \Lambda)$ have well defined limiting distributions. Next, consider

$$\sqrt{T}[Q'(\hat{M} - M(0))Q] = \sqrt{T}(\tilde{M}^* - \Lambda), \quad \tilde{M}^* = Q'\tilde{Q}\tilde{\Lambda}\tilde{Q}'Q, \tag{30}$$

and put

$$\sqrt{T}(\tilde{Q}'Q - I_q) = C, \quad \sqrt{T}(\tilde{\Lambda} - \Lambda) = D, \quad \sqrt{T}(\tilde{M}^* - \Lambda) = G. \tag{31}$$

Note that by Theorem 1, C, D, G, are all a.c. finite random variables, in the sense that they have well defined limiting distributions, and thus may assume the values $\pm\infty$ only on a set of measure zero. It follows immediately that

$$\sqrt{T}(\tilde{M}^* - \Lambda) = G \sim D + \Lambda C - C\Lambda. \tag{32}$$

Since

$$g_{ii} = d_{ii}, \quad g_{ij} = (\lambda_i - \lambda_j)c_{ij}, \quad \text{or} \quad c_{ij} = g_{ij}/(\lambda_i - \lambda_j), \quad i \neq j, \tag{33}$$

this concludes the proof of the theorem, for the case where all characteristic roots are different so that the elements c_{ij} are well defined. We now examine the case where

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \quad \Lambda_2 = 0,$$

and Λ_1 is a diagonal matrix containing the ($r = q - r_0$) positive roots under the null of cointegration; evidently, Λ_2 contains the r_0 zero roots. Partitioning the other matrices conformably we determine

$$\sqrt{T}(\tilde{M}^* - \Lambda) = G \sim \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} + \begin{bmatrix} \Lambda_1 C_{11} - C_{11}\Lambda_1 & \Lambda_1 C_{12} \\ -C_{21}\Lambda_1 & 0 \end{bmatrix}. \tag{34}$$

Consequently, we have again

$$\begin{aligned} g_{ii} &= d_{ii}, \quad i = 1, 2, \ldots, q, \quad g_{ij} = (\lambda_i - \lambda_j)c_{ij}, \quad i, j = 1, 2, \ldots, r, \quad i \neq j; \\ g_{ij} &= \lambda_i c_{ij}, \quad i = 1, 2, \ldots, r, \quad j = r+1, r+2, \ldots, q; \\ &= \lambda_j c_{ij}, \quad i = r+1, r+2, \ldots, q, \quad i = 1, 2, \ldots, r; \\ &= 0, \quad i, j = r+1, r+2, \ldots, q, \quad i \neq j. \end{aligned} \tag{35}$$

From the construction above it follows that if we partition G conformably

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix},$$

we must conclude that $G_{22} = D_{22}$ and, moreover that $D_{22} \xrightarrow{\mathrm{d}} 0$. To verify this claim we note that, using the construction of Eq. (30), the covariance matrix of the limiting distribution of G_{22} may be obtained from the result

$$\sqrt{T}(Q' \otimes Q')\mathrm{vec}[\hat{M} - M(0)] \xrightarrow{\mathrm{d}} N(0, \Psi), \quad \Psi = (Q' \otimes Q')\Psi^*(Q \otimes Q),$$

where

$$\Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix} = \begin{bmatrix} [Q_1' \otimes Q']\Psi^*[Q \otimes Q_1] & [Q_2' \otimes Q']\Psi^*[Q \otimes Q_2] \\ \\ [Q_1' \otimes Q']\Psi^*[Q \otimes Q_2] & [Q_2' \otimes Q']\Psi^*[Q \otimes Q_2] \end{bmatrix}. \tag{36}$$

Thus, the **marginal** limiting distribution of G_{22} obeys

$$G_{22} \xrightarrow{\mathrm{d}} N(0, \Psi_{22}),$$

or what is equivalent

$$\sqrt{T}(\hat{\Lambda}_2 - \Lambda_2) \xrightarrow{\mathrm{d}} N(0, \Psi_{22}), \quad \Psi_{22} = [Q_2' \otimes Q']\Psi^*[Q \otimes Q_2]. \tag{37}$$

Since $\Psi_{22} = 0$, we have that

$$\sqrt{T}(\tilde{\Lambda}_2 - \Lambda_2) \xrightarrow{\mathrm{d}} 0, \quad \text{and hence that} \quad \sqrt{T}(\tilde{\Lambda}_2 - \Lambda_2) \xrightarrow{\mathrm{P}} 0.$$

q.e.d.

Remark 3. Since

$$\sqrt{T}\mathrm{tr}\tilde{\Lambda}_2 \quad \text{and} \quad \hat{\Psi}_{22} = [\hat{Q}_2' \otimes \hat{Q}_2']\hat{\Psi}^*[\hat{Q}_2 \otimes \hat{Q}_2]$$

converge at the same rate to zero we have, as a very good approximation, that

$$\sqrt{T}\mathrm{tr}\tilde{\Lambda}_2 \sim N_{tr}(0, \phi^2), \quad \phi^2 = e'\Psi_{22}e, \tag{38}$$

where e is a suitably dimensioned vector of unities and $N_{tr}(0, \phi^2)$ is the normal **truncated at zero**. We may approximate the finite sample distribution of $\mathrm{tr}\tilde{\Lambda}_2 \sim N_{tr}(0, \hat{\phi}_T^2)$, where $\hat{\phi}_T^2 = e'\hat{\Psi}_{22}e/T$. Next, renormalize it so that its integral over $(0, \infty)$ is unity and obtain its mean, say μ_T. The test of the null that $\mathrm{tr}\,\Lambda_2 = 0$ may thus be carried out through the test statistic

$$\mathrm{CCT}(\mathrm{r}_0) : \frac{\sqrt{T}(\mathrm{tr}\tilde{\Lambda}_2 - \mu_T)}{\hat{\phi}} \xrightarrow{\mathrm{d}} N(0,1), \quad \hat{\phi}^2 = e'\hat{\Psi}_{22}e, \tag{39}$$

which is seen to be a standard test.

In the framework of the conformity test for cointegration, the presence of zero roots is almost self evident by visual inspection. The separation of roots is striking, especially when compared to the separation of characteristic roots in the context of the Likelihood Ratio test framework given in Johansen (1988), (1991). This will be clearly illustrated in the limited Monte Carlo experiment below.

Corollary 2. The characteristic roots of $\hat{M}$, and hence their limiting distributions, are exactly those of $\tilde{M}^*$.

Proof: Obvious since Q is a fixed orthogonal matrix.

Corollary 3. The distribution of the (associated) characteristic vectors is given by the distribution of QC'.

Corollary 4. A test on the rank of cointegration may be carried out as follows: let $\hat{\lambda}_{(i)},\ i = 1, 2, \ldots, q$ be the characteristic roots of

$$|\lambda I_q - \hat{M}| = 0,$$

arranged in **decreasing order**, and let it be desired to test the hypothesis

H_0: the rank of cointegration is r

as against the alternative

H_1: the rank of cointegration is $r + s \leq q$.

Consider the entity

$$\tau^* = \frac{\sum_{i=1}^{s} \hat{\lambda}_{(r+i)}}{\sqrt{e'\hat{\Psi}_{22(s)}e}},$$

where e is a suitably dimensioned vector of unities, and $\hat{\Psi}_{22(s)}$ is the estimated covariance matrix of the limiting distribution of the s roots in question. Thus, under H_0

$$\tau^* \sim N_{tr}(0, 1). \tag{40}$$

The remaining problem is to determine the matrix Ψ_q corresponding to the covariance matrix of the limiting distribution of the roots. Denoting the elements of the $q^2 \times q^2$ matrix Ψ by (ψ_{ij}), $i, j = 1, 2, \ldots, q^2$, we give below the elements in the upper triangular position of the matrix Ψ_q.

$$\Psi_q = \begin{bmatrix} \psi_{11} & \psi_{1,q+2} & \psi_{1,2q+3} & \psi_{1,3q+4} & \cdots & \psi_{1,q^2} \\ * & \psi_{q+2,q+2} & \psi_{q+2,2q+3} & \psi_{q+2,3q+4} & \cdots & \psi_{q+2,q^2} \\ * & * & \psi_{2q+3,2q+3} & \psi_{2q+3,3q+4} & \cdots & \psi_{2q+3,q^2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & * & * & \psi_{q^2,q^2} \end{bmatrix} \tag{41}$$

Remark 4. The major computational burden entailed by this procedure is the estimation of the spectral matrix of the vector process $\hat{z}'_{t-1\cdot} \otimes \hat{z}'_{t-1\cdot}$, where $\hat{z}_{t\cdot} = X_{t-1\cdot}\hat{\Pi}(1)$ is the estimated cointegral vector; or, equivalently, the estimation of the covariance matrix, Ψ, of the limiting distribution above.

3 Conformity Cointegration Tests for $MIMA(k)$ Processes

In this section we shall show that the results above are almost identically applicable, even if we assume that the X-process is $MIMA(k)$, $k < \infty$, i.e.

$$(I - L)X_{t\cdot} = \epsilon_{t\cdot}A(L), \quad A(L) = \sum_{j=0}^{k} A_j L^j, \quad A_0 = I_q, \tag{42}$$

and $|A(z)| = 0$ has r_0 **unit roots**. We begin by noting that in Proposition 2a in Dhrymes (1995) Chapter 7, it is shown that the X-process has the reprsentation

$$X_{t\cdot}H(L) = \epsilon_{t\cdot}a(L)I_q, \tag{43}$$

where $H(L)$ is a matrix polynomial lag operator, and $a(L)$ is an invertible scalar polynomial lag operator, respectively, of degrees

$$m_1 = (q-1)(k-1) + r_0, \quad m_2 = q(k-1) + r_0. \tag{44}$$

There are two differences between this formulation and the $VAR(n)$ formulation. First, $MIMA(k)$ is **equivalent to an** $ARMA(m_1, m_2)$ process. Thus, the properties of the error term in Eq. (43) are considerably more complex than in the $VAR(n)$ formulation. Second, in the $MIMA(k)$ formultion the parameter k need not be specified explicitly; but even for moderate values like $k = 10$ and $q = 6$, the autoregressive component of the resulting ARMA process is of order $(q-1)(k-1) + r_0 = 54 + r_0$! As we shall show below, however, an accurate specification of the lags is not essential for establishing the limiting distribution of the conformity cointegration test statistics. Thus, Theorems 1, 2, 3 and Corollaries 1, through 4, remain valid if

$$(I - L)X'_{t\cdot} = \sum_{j=0}^{k} A_j \epsilon'_{t-j\cdot}, \tag{45}$$

the ϵ-process is $MWN(\Sigma)$, and is required to obey a condition analogous to that in Eq. (12). We now prove

Theorem 3. Consider the $MIMA(k)$ model of Eq. (41) and suppose the moment condition in Eq. (12) holds, where the cointegral vector is defined by

$z_{t\cdot} = X_{t\cdot}H(1)$, the entity $H(1)$ being as implicitly defined in Eq. (43). Then the conclusions of Theorems 1, 2 and Corollaries 1 through 4 remain valid, whether or not the ARMA representation of the X-process in Eq. (43) is correctly specified.

Proof: From Eq. (43) we may obtain the equivalent representation

$$(I-L)X_{t\cdot} = -X_{t-1\cdot}H(1) + \sum_{j=1}^{m_2-1}(I-L)X_{t-j\cdot}H_j^* + u_{t\cdot}, \tag{46}$$

$$H(L) = H(1) - (I-L)H^*(L), \quad H_j^* = \sum_{i=j+1}^{m_2} H_i.$$

To maintain maximal correspondence with the preceding discussion, let

$$X_1 = (\Delta X_{t-1\cdot}, \Delta X_{t-2\cdot}, \ldots, \Delta X_{t-n+1\cdot}), \tag{47}$$

$$X_2 = (\Delta X_{t-n\cdot}, \Delta X_{t-n-1\cdot}, \ldots, \Delta X_{t-m_2+1}),$$

and, in the obvious notation write the observations on this model as

$$\Delta P = -P_{-1}H(1) + X_1 H_{(1)}^* + X_2 H_{(2)}^* + U, \quad U = (u_{t\cdot}), \quad u_{t\cdot} = a(L)\epsilon_{t\cdot}. \tag{48}$$

Note that since $a(L)$ is invertible, the u-process is strictly stationary. Suppose in the estimation of $H(1)$ **we neglect** the variables in X_2, thus running a regression of ΔP on (P_{-1}, X_1). The least squares estimator of $H(1)$ is given by

$$\hat{H}(1) = -H(1) + (V'V)^{-1}V'U + (V'V)^{-1}V'X_2 H_{(2)}^*. \tag{49}$$

Moreover, we note that for any $\alpha \in [0, 1)$

$$T^{\alpha}\hat{H}(1) \xrightarrow{\text{P}} -H(1). \tag{50}$$

If we now define the matrix

$$\hat{M} = \frac{1}{T}\hat{H}(1)' P_{-1}' P_{-1}\hat{H}(1), \tag{51}$$

we have, by precisely the same arguments employed earlier, that

$$\sqrt{T}(\hat{M} - M_{zz}) \sim \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(z_{t-1\cdot}' z_{t-1\cdot} - M_{zz}), \tag{52}$$

where now

$$z_{t\cdot} = X_{t\cdot}H(1), \quad \frac{H(1)' P_{-1}' P_{-1} H(1)}{T} \xrightarrow{\text{a.c.}} M_{zz} \geq 0. \tag{53}$$

In addition, note that in deriving the limiting distribution of the entity in Eq. (9) **we did not make use of the fact that** the ϵ-process was one of i.i.d. random variables. We only made use of the fact that it was a **strictly stationary process.** Since the same condition holds in this case as well, the conlcusions of Theorems 1, 2, 3, and Corollaries 1 through 4, continue to be valid.

q.e.d.

Before we proceed to the emirical implementation of the procedures above it is useful to explore two aspects: first, what is the intuitive content of the derivation of the limiting distribution of the matrix $\hat{M}$ and second, what is the connection between the roots of $\hat{M}$ and the roots that appear in the LR procedure (Johansen). We develop these issues in the two remarks below.

Remark 5. Since

$$\hat{\Pi}(1) = (V'V)^{-1}V'W, \quad W = N\Delta P,$$

we might as well have written

$$\hat{M} = \frac{1}{T}W'V(V'V)^{-1}P'_{-1}P_{-1}(V'V)^{-1}V'W.$$

Doing so, leads us to observe that since

$$(V'V)^{-1}P'_{-1}P_{-1} \xrightarrow{\mathrm{d}} I_q, \quad \text{and thus} \quad (V'V)^{-1}P'_{-1}P_{-1} \xrightarrow{\mathrm{P}} I_q,$$

we have

$$\hat{M} \sim \frac{1}{T}W'V(V'V)^{-1}V'W. \tag{54}$$

If the representation above were appropriate, it would imply that the theory developed in this paper can equally well be developed on the basis of the entity

$$\hat{M}^{(1)} = \frac{1}{T}\hat{\Pi}(1)'V'V\hat{\Pi}(1) = \frac{1}{T}W'V(V'V)^{-1}V'W. \tag{55}$$

This is, however, incorrect since $\hat{M}^{(1)}$ need not have the same probability limit as $\hat{M}$, and need not have the same limiting distribution. To see this, expand its representation to obtain

$$\begin{aligned}
\hat{M}^{(1)} &= \frac{1}{T}(B_{11} + B_{12} + B'_{12} + B_{22}) \\
B_{11} &= \Pi(1)'P'_{-1}V(V'V)^{-1}V'P_{-1}\Pi(1), \quad B_{12} = \Pi(1)'P'_{-1}V(V'V)^{-1}V'U \\
B_{22} &= U'V(V'V)^{-1}V'U
\end{aligned}$$

It is clear that

$$\frac{1}{\sqrt{T}}B_{22} \xrightarrow{\mathrm{P}} 0, \quad \frac{1}{\sqrt{T}}B_{12} \xrightarrow{\mathrm{P}} 0,$$

and moreover

$$\frac{1}{T}B_{11} = \frac{1}{T}\Pi(1)'P_{-1}'P_{-1}\Pi(1) - \frac{\Pi(1)'P_{-1}'X}{T}\left(\frac{X'X}{T}\right)^{-1}\frac{X'P_{-1}\Pi(1)}{T}.$$

Eivdently,

$$\frac{1}{T}B_{11} \xrightarrow{\mathrm{ac}} M_{zz}^{*} = M_{zz} - M_{zx}M_{xx}^{-1}M_{xz}. \tag{56}$$

Thus, M_{zz}^{*} is the **conditional covariance matrix** of the cointegral vector $z_{t\cdot} = X_{t-1\cdot}\Pi(1)$, conditioned on the σ-field induced by the elements of the matrix X, i.e. the differences $\Delta X_{t-i\cdot}$, $i = 1, 2, \ldots, n-1$. Moreover, the limiting distribution is obtained from

$$\sqrt{T}(\hat{M}^{(1)} - M_{zz}^{*}) \sim \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(z_{t-1\cdot}'z_{t-1\cdot} - M_{zz}) \tag{57}$$

$$-M_{zx}M_{xx}^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[x_{t\cdot}'z_{t-1\cdot} - M_{xz}]\right).$$

Evidently, the behavior of $\hat{M}$ and $\hat{M}^{(1)}$ will be **identical if** $M_{zx} = 0$!

Remark 6. We now turn to the connection between the (characteristic) roots encountered in the conformity test, and those encountered in the LR test. If we concentrate the likelihood function, as in Johansen (1988), (1991), we ultimately find that we need to **minimize** with respect to B, which is a $q \times r$ matrix of rank r, the determinant

$$D(B) = |W'W - W'VB(B'V'VB)^{-1}B'V'W|. \tag{58}$$

After considerable manipulation, we determine that this requires us to obtain the (characteristic) roots and vectors of

$$|\lambda V'V - V'W(W'W)^{-1}W'V| = 0, \tag{59}$$

The solution to the problem is to select the r largest roots, and their associated characteristic vectors. The latter serve as the estimator of the matrix B in the rank factorization $\Pi(1) = B\Gamma'$. Note further that, under the null of cointegration, the remaining roots are zero. Hence, in the LR procedure the rank of cointegration is simply the number of positive roots in the limit version of Eq. (59). In the conformity test context, the rank of cointegration is determined by the positive roots of the limit of $\hat{M}$. An important question then is: how are the roots of Eq. (59) related to the roots of $\hat{M}$? Using a number of properties of determinants, particularly Proposition 43 in Dhrymes (1984), p. 51, we have that the characteristic roots of Eq. (59) are precisely those of

$$|\lambda W'W - W'V(V'V)^{-1}V'W| = 0. \tag{60}$$

Thus, basically, the LR (Johansen) procedure determines the rank of cointegration by the number of positive roots of the limit of $\hat{M}^{(1)}$, **in the metric** of $M^*_{x_0x_0}$. The latter is simply the conditional covariance matrix of $\Delta X_{t\cdot}$, conditioned on the same σ-field as above. It is this feature that renders such roots **less than unity**, and thus impedes the effective separation of roots in empirical applications. In contrast, the conformity approach determines the rank of cointegration by the positive characteristic roots of the limit of $\hat{M}$, which is the **unconditional covariance** matrix of the cointegral vector, **in the metric of the identity matrix**. Thus, the roots are not compressed by measuring them in terms of "units" of a possibly large covariance matrix, and this contributes to a very effective separation of the zero roots in an empirical context.

4 Empirical Implementation and Monte Carlo Results

4.1 Estimation of the Asymptotic Covariance Matrix

The empirical implementation of conformity cointegration tests entails the following steps.

i. Obtain **unrestricted** estimators of $\Pi(1)$, or $H(1)$, by regressing $\Delta X_{t\cdot}$ on $X_{t-1\cdot}$ and a number of lag differences, say $\Delta X_{t-1\cdot}$ through $\Delta_{t-n+1\cdot}$.

ii. Form the matrix $\hat{M} = (1/T)\hat{\Pi}(1)'P'_{-1}P_{-1}\hat{\Pi}(1)$ and obtain its characteristic roots (and vectors).

iii. Estimate the covariance matrix of the limiting distribution of the (centered) characteristic roots.

iv. Carry out the tests.

All steps, except iii, have been dealt with in earlier sections. Here we discuss the estimation of the covariance matrix. The simplest approach is to determine the appropriate estimator of the matrix Ψ^* exhibited in Eq. (20), or of Ψ, as exhibited in Eq. (36). Since the estimator involves an infinite series we require a covariance window, or a **kernel**. For this purpose we choose the quadratic spectral kernel given by

$$\kappa(\tau) = \frac{25}{12\pi^2\tau^2}\left(\frac{\sin(6\pi\tau/5)}{(6\pi\tau/5)} - \cos(6\pi\tau/5)\right). \tag{61}$$

This kernel[4] was shown to be optimal in the estimation of multivariate probability functions, see Epanechnikov (1969), and has been employed in the literature of econometrics by Andrews (1991). It has the property, shown in Andrews, that it ensures the that the estimated covariance matrices are **positive semidefinite**. Thus, the estimated covariance matrix is given by

$$\begin{aligned}
\hat{\Psi}_T^* &= \frac{1}{T}\left(\sum_{t=1}^{T}\hat{\zeta}_{t\cdot}'\hat{\zeta}_{t\cdot}\right) + \sum_{\tau=1}^{T-1}\kappa\left(\frac{\tau}{b}\right)(C(\tau) + C(\tau)'),\\
\hat{z}_{t\cdot} &= X_{t\cdot}\hat{\Pi}(1), \quad \hat{\zeta}_{t\cdot} = \hat{z}_{t-1\cdot}\otimes\hat{z}_{t-1\cdot} - m(0)',\\
C(\tau) &= \frac{1}{T}\sum_{t=1}^{T-\tau}\hat{\zeta}_{t\cdot}'\hat{\zeta}_{t+\tau\cdot}.
\end{aligned} \tag{62}$$

The parameter b is the data determined automatic bandwidth which ranges, in the Monte Carlo results below, from 2.5 to 4.3.

4.2 Monte Carlo Results

4.2.1 Histograms and Empirical Densities

Various aspects of the developments above have been investigated by means of a limited Monte Carlo study.

The purpose of the presentation here is not to give extensive results, but rather to illustrate some of the properties of this procedure, and compare its results to those that are obtained by LR tests, as in Johansen (1988), (1991), and the determination of "stochastic trends", as in Stock and Watson (1988).

[4] Note that since

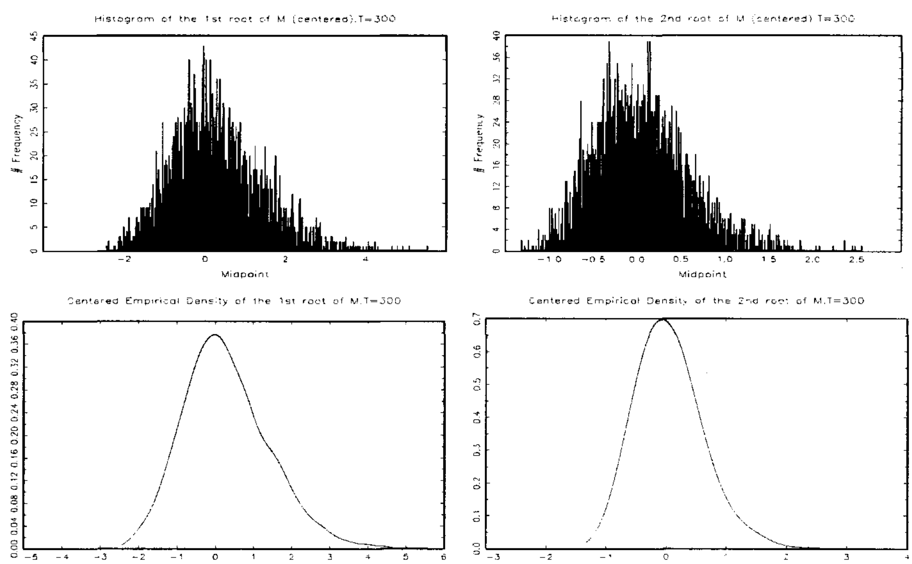
$$\begin{aligned}
\frac{\sin x}{x} &= 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!}, \quad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots, \quad \text{we obtain}\\
\frac{\sin x}{x} - \cos x &= \frac{2x^2}{3!} - \frac{4x^4}{5!} + \frac{6x^6}{7!} - \ldots. \quad \text{Hence,}\\
\lim_{\tau\to 0}\kappa(\tau) &= 1, \quad \text{and, thus, } \kappa \text{ is } \textbf{continuous} \text{ at zero.}
\end{aligned}$$

Finally, we apply this procedure to the (extended) money demand data set investigated in Stock and Watson (1993) and compare the conclusions of the three procedures.

The Monte Carlo model is a $VAR(3)$, with three variables. The underlying stochastic sequence $\{X_{t\cdot} : t \geq 1\}$ is generated recursively from a sequence of i.i.d. vectors $\{\epsilon_{t\cdot} : t \geq 1\}$, which are $N(0, \Sigma)$, $\Sigma > 0$. The elements of Σ exhibit only moderate correlation among the components of the ϵ-sequence.[5] We present results for two parametric configurations. In both there is **only one unit root**. In the first, (SMSR), the stationary roots are small (hence the acronym) and they range (in absolute value) between .1 and .35; in the second, (LGSR), the stationary roots range between .27 and .95. We have carried out two sets of 3,000 replications, one with sample size 100, and one with sample size 300.

Graph 1

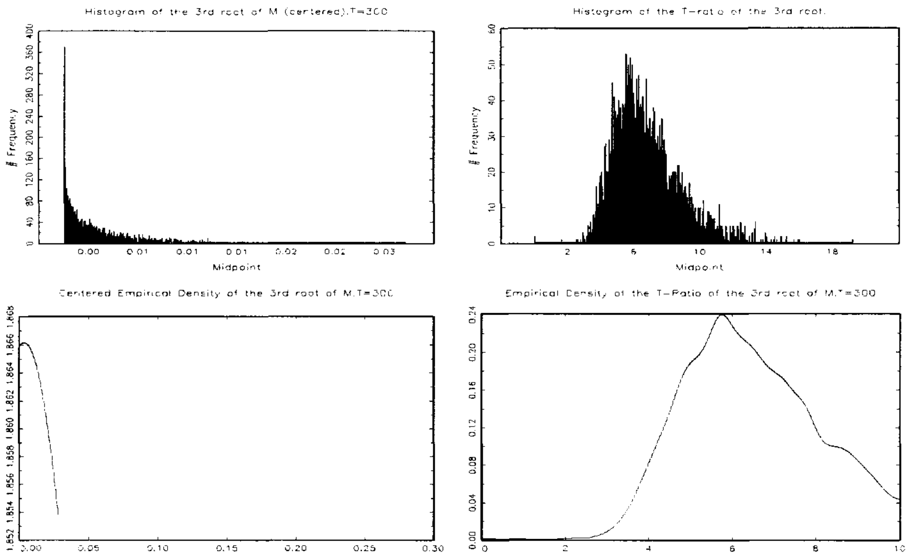
Historgrams and Empirical Densities of $(\hat{\lambda}_i - \lambda_i)$, $i = 1, 2$.



[5] In some Monte Carlo studies of cointegration, investigators discard a number of initial observations in order "to eliminate start-up effects". We have not followed this procedure since in most such studies (including this) it is either implicitly or explicitly assumed that $X_{-s\cdot} = 0$, for $s \geq 0$. Such an assumption is incompatible with the practice of discarding initial observations.

Graph 2

Historgram and Empirical Density of $\hat{\lambda}_{\min}$ and of $\sqrt{T}\hat{\lambda}_{\min}/\hat{s}$.



In Graphs 1 and 2 above, we illustrate the behavior of the (characteristic) roots of the matrix $\hat{M}$. The histograms and associated empirical densities are centered about the true roots of $M(0)$, respectively, 5.739, 2.406, and zero. The results are identical for sample size 100 and are not presented.

In deriving the empirical densities we have employed a normal kernel with bandwidth $1.06N^{-.2}$, where N is the number of replications. This is optimal, according to Silverman (1986). In our case, with 3,000 replications the bandwidth is $b = .2137$.

In Graph 1, we give the historgram and empirical density of the estimated nonzero roots, centered about the true roots. Notice that the distributions are almost symmetric about the true roots, although as we shall see in Table 1 below, there is still some bias. None the less it is quite apparent that we have fairly rapid convergence to the (normal) limiting distribution. In Graph 2, we give the histogram and empirical density of $\hat{\lambda}_{\min}$, i.e. the estimator of the zero root, and its "t-ratio "$(\sqrt{T}\hat{\lambda}_{\min}/s)$, where s is the estimated standard error. The histogram has a rather lean tail, and illustrates well the truncated normal character of the third root of $\hat{M}$, corresponding to the zero root of $M(0)$. This is also reflected in the empirical density.

The shape of the empirical density of $(\sqrt{T}\hat{\lambda}_{\min}/s)$, is of no particular significance in this context, but it does illustrate the fact that although both numerator and denominator tend to zero their ratio exhibits a certain (distributional) stability. The historgram and empirical density of the statistic $\sqrt{T}(\hat{\lambda}_{\min} - \hat{\mu}_T)/s$, not shown, exhibits an approximate unit normal shape, but it is not centered precisely at zero.

4.2.2 Separation of Zero Roots

In Table 1, below, we present an illustration of the phenomenon alluded to earlier, viz. that the separation of roots is much more striking in the conformity test than in the LR test context.

Table 1

Root Separation, Conformity and Johansen (LR) Tests
VAR(3), 3,000 Replications

	Test Type			
Experiment	**Conformity**			**Johansen (LR)**
	Mean Root	True Root	Bias	Mean Root
SMSR, T = 100				
λ_1	6.609	5.739	0.910	.458
λ_2	2.456	2.406	0.050	.258
λ_3	0.010	0.000	0.010	.012
SMSR, T = 300				
λ_1	6.039	5.739	0.300	.443
λ_2	2.436	2.406	0.030	.247
λ_3	0.003	0.000	0.003	.004
LGSR, T = 100				
λ_1	681.711	709.528	-27.817	0.852
λ_2	10.899	8.989	1.910	0.283
λ_3	0.060	0.000	0.060	0.015
LGSR, T = 300				
λ_1	698.108	709.528	-11.420	0.849
λ_2	9.481	8.989	0.492	0.258
λ_3	0.008	0.000	0.008	0.005

Note that even for a sample of size 100, the zero root is very effectively separated in the conformity context. For the SMSR case, the second root is 245.6 times the third root, and the first root is 2.69 times the second root; in the LR (Johansen) context the second root is 21.5 times the third root, and the first root is 1.77 times the second root. For the LGSR case in the conformity context, the second root is 181.65 times the third root, and the first root is 62.55 times the second root; in the LR (Johansen) context, the corresponding magnitudes are, respectively, 18.87 and 3.01.

For sample size 300, in the SMSR case the second root is 812 times the third root, and the first root is 2.48 times the second root [6] in the conformity context; in the LR (Johansen) case the corresponding magnitudes are 61.75 and 1.79. For the LGSR case, in the conformity context, the ratio of the second to the third root is 1185.12, and the ratio of the first to the second root is 73.63, which is close to the ratio of the true corresponding roots, 78.93; in the LR (Johansen) case the corresponding magnitudes are 51.6 and 3.29.

While the relation between nonzero roots is not of particular significance, since it evidently depends on the structure of the model, the relation between the estimator of a zero root, and the estimator of the smallest nonzero root **is of particular significance**. In the conformity context, even for moderate sample size, the existence of a zero root is obvious by visual inspection. This is a highly desirable property, since in empirical applications we cannot be certain of the true underlying model; thus, a test procedure that so effectively separates the zero roots is of considerable significance.

Finally, note the appreciable reduction in the bias of the finite sample distribution of the estimated roots.

Table 2, below, [7] presents results for certain tests on the models considered in Table 1, which involve a **single unit root**, as well as results from a model characterized by two unit roots and large stationary roots, ranging in absolute value from .6, to .95. It shows that the conformity cointegration test statistic, testing the hypothesis that $\mathrm{tr}\,(M) > 0$, performs quite well; when the standard test, however, is applied to the estimator of a zero root, the (rank test) statistic $\sqrt{T}\hat{\lambda}_{\min}/s$, performs rather poorly in testing the hypothesis that $\lambda_3 = 0$ and, *mutatis mutandis*, in testing the hypothesis $\lambda_2 + \lambda_3 = 0$, in the case of two unit roots. On the other hand, the test of such hypothesis based on the truncated normal distribution performs rather well, using the centering parameter ("bias" correction)

$$\hat{\mu}_T = \left(\frac{(\pi s^2/2)}{T^\alpha}\right)^{1/2}, \quad \text{with} \quad \alpha = \frac{r_{\max} - r_{\min}}{3},$$

where $r_{\max}$, $r_{\min}$ are (in absolute value) the largest and smallest stationary roots of the estimated companion matrix. Its major deficiency is that its empirical size is somewhat smaller that the specified size, ranging from .01 to .05.

The LR (Johansen) test performs generally well although in the case of large stationary roots, its empirical size exceeds appreciably the specified size (.05). For sample size 100 (and one unit root), the empirical size of the test of $\lambda_3 = 0$

[6] Note that in the conformity context the true roots are 5.739, 2.406 and 0; thus the ratio between the first and second **true roots** is 2.38, which is reasonbly close to what we obtain in samples of size 100 and 300.

[7] In Table 2, the heading $\lambda_{2,3} = 0$ stands for the null hypothesis $\lambda_2 + \lambda_3 = 0$.

is .11; for two unit roots and sample size 100, the empirical size of the test of $\lambda_{2,3} = \lambda_2 + \lambda_3 = 0$ is .15. Performance improves as the sample increases; for 300, the corresponding empirical size is .06 and .09.

For the S & W test, we have obtained the three principal compoments of the variables of the VAR, "filtered the data", and applied a VAR(1) model. The tests reported are based on the first, second, and third roots of the estimated "autocovariance" matrix. Thus, for example, the entry under 3 UR is the frequency of rejection of the hypothesis that the third root is unity; the entry under 2 UR is the frequency of rejection of the hypothesis that the second root is unity. The results often indicate stationarity, as implied by rejection of the hypothesis that the first root is unity (not given in Table 2). For example, in the two unit root case with large stationary roots, sample size 100, the hypothesis that the first root is unity is **rejected** in 35% of the 3,000 replications. When the sample size is 300, it is **rejected** in 54% of the replications. Thus, the procedure erroneously imputes stationarity at an inordinately high frequency.

Table 2

Performance of Conformity, Johansen (LR), and S & W Tests
VAR(3), 3,000 Replications. Size of Test is .05

TEST TYPES AND FREQUENCY OF REJECTION							
Standard Conformity and Johansen				Using N_{tr}		S & W	
$H_0 \longrightarrow$	tr $M > 0$	$\lambda_{2,3} = 0$	$\lambda_3 = 0$	$\lambda_{2,3} = 0$	$\lambda_3 = 0$	3UR	2 UR
Test Type							
SMSR	T = 100	1 UR					
Conformity	0	1	0.72	n.a.	.05	n.a.	n.a.
Johansen	n.a.	.99	.06	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	1	.97
SMSR	T = 300	1 UR					
Conformity	0	1	.55	n.a.	.02	n.a.	n.a.
Johansen	n.a.	1	.06	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	1	1
LGSR	T = 100	1 UR					
Conformity	.01	.99	.76	n.a.	.02	n.a.	
Johansen	n.a.	1	.11	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	.003	.83
LGSR	T = 300	1 UR					
Conformity	0	1	.65	n.a.	.01	n.a.	n.a.
Johansen	n.a.	1	.06	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	.003	.99
LGSR	T = 100	2 UR					
Conformity	0	.998	.950	.054	.009	n.a.	n.a.
Johansen	n.a.	.158	.029	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	.0157	.652
LGSR	T = 300	2 UR					
Conformity	0	.999	.997	.001	.001	n.a.	n.a.
Johansen	n.a.	.094	.015	n.a.	n.a.	n.a.	n.a.
S & W	n.a.	n.a.	n.a.	n.a.	n.a.	0	.166

4.3 Application to Money Demand in the US

Here we apply the procedures developed above to the annual data given in Stock and Watson (1993). The data set was extended to 1995. The objective is to illustrate the application of various tests on the presence, and rank, of cointegration. It is not our intention to offer substantive comment on the nature of the demand for money in the US over the twentieth century, or its stability. The data comprise the log of the money supply (M_1), denoted by m , the log of the implicit deflator of the net national product, denoted by p , the log of the (real) net national product, denoted by y , and the (annualized) interest rate on six month commercial paper (in per cent) denoted by r . Thus, the vector to be investigated is $X_{t\cdot} = (m_t - p_t, y_t, r_t)$; we treat the entire period as parameter-homogeneous, in view of the S & W finding of stability over this period. Placing the problem in a trivariate $VAR(3)$ context, one has the following results:

i. In terms of the conformity cointegration test framework, the characteristic roots of

$$\hat{M} = \frac{1}{T}\hat{\Pi}(1)' P_{-1}' P_{-1} \hat{\Pi}(1),$$

are .19902, 9.336 $\times 10^{-5}$, and 9.580 $\times 10^{-6}$. The ratio of the second to the third root is about 9.7 and the ratio of the first to the second is about 2131. In view of the prior experience with $VAR(3)$ systems, we may be inclined to conclude that there are **two** zero roots, and thus **two unit** roots [8] of the charactreristic equation $|\Pi(z)| = 0$. A formal test for $\mathrm{tr}\,(M) = .001$ yields the test statistic 2.526, and thus acceptance of the hypothesis that $\mathrm{tr}\,(M) > 0$ at the 5% as well the 1% level of significance.

Based on the bivariate normal truncated at zero, the formal test that the sum of the last two roots is zero yields the test statistic - 1.286, which results in acceptance. Since, in the conformity context, we have information on the (limiting) distribution of both the zero **and** the nonzero roots, we may also test the same hypothesis, on the assumption that the sum of the last two roots is **positive**, which may be based on the limiting distribution for positive roots developed above. In this case we obtain the test statistic 2.869 which supports the view that the second root is positive. This is, of course, an inappropriate test if, in fact, the last two roots are null, as suggested by the magnitude of the estimated (last two) roots of $\hat{M}$. [9] A test that the third root is zero, based on the normal truncated at zero,

[8] Of course, considering the rather small size of the first root, one might equally well draw the conclusion that a $VAR(3)$ is **not** an appropriate specification for this relation, and that the system is, in fact, stationary.

[9] Notice that this amounts to testing a null hypothesis based on the distribution of the test statistic under the alternative. As such it is really not a test of significance, but rather an attempt to explore power issues.

yields the test statistic -1.093, which results in acceptance. Thus, in the context of the conformity cointegration test, we find that this system is **cointegrated of rank** 1.[10]

Finally, the characteristic roots of the estimated companion matrix are (in absolute value): .9845, .9636, .6381, .6381, .5438, .5438, .2742, .2742, and .0848. We note that, as estimated, the system has three pairs of complex roots, and two real roots near unity; all real roots are positive, except the last which is negative.

ii. In the LR (Johansen) context, one finds the characteristic roots .1624, .0289, and .0032; the hypothesis that the cointegration rank is 1, is equivalent to the hypothesis that the sum of last two roots is zero; since the test statistic is $-95\ln(1-.0321)=3.098$ and, at the 5% level of signficance, the critical value is 12.53 the hypothesis is accepted. Note that in this context, it is not possible to test the hypothesis that the sum of the last two roots is zero, based on the distribution of positive roots. The latter may be derived, but is not explicitly available in the literature. Quite likely, in view of the size of the LR (Johansen) roots, the conclusion would be the same as in the conformity test. At any rate such considerations address the issue of the power of the test(s), which is not the focus of this exercise.

iii. In the S & W "filtering the data" context, we first extract the principal components of $X_{t\cdot}$; in obtaining principal components one subtracts the sample mean from each observation, thus dealing with centered data.[11] The characteristic roots of the resulting "sample covariance" matrix are 8.2067, 1.0267, and .0115. We note that the ratio of the largest to the smallest root is about 713 and, indeed, the principal component corresponding to the largest root accounts for about 89% of the variability of the data; the principal component corresponding to the second largest root, accouts for about 11% of the variation in the data. This is an even more extreme case than that reported in Stone (1947) – see also Dhrymes (1970), p.64.

We recall that in the S & W procedure one (subsequently) obtains a certain filtered vector[12] say $\xi_{t\cdot}$, regresses it on $\xi_{t-1\cdot}$, thus obtaining an estimator, $\hat{A}$, of the "first order autocorrelation matrix". If $\hat{\mu}_i$, $i=1,2,3$, are the (ordered) characteristic roots of $\hat{A}$, the S & W procedure tests the

[10] We should note here that, to the extent that the tests are **sequential**, the size of the test is overstated.

[11] In the standard (stationary) context the resulting entity is an estimator (under normality, a ML estimator) of the underlying covariance matrix.

[12] In keeping with the basic specification of the model, we have used three lags in the filtering phase.

hypotheses $\mu_i = 1$. Since the roots are in (decreasing) order, to test for stationarity we need only test that $\mu_1 = 1$. If we reject this hypothesis, thereby accepting that $\mu_1 < 1$, we conclude that the system is **stationary**. If not, we proceed to determine the number of unit roots, and thus the rank of cointegration, or absence thereof. The characteristic roots in question are .9892, .9139, .2288, so that we have **three real roots**, all less than unity in absolute value. The S & W test statistic for the first root is -.9906, that for the second root is - 7.919, and for the third it is - 70.947. From Table 1 of Stock and Watson (1988) we find that, at the 5% level of significance, the critical value for the first root is - 2.53. Thus, the hypothesis that the largest root is unity is **accepted**; the critical value for the second, is - 11.1; thus this hypothesis is also **accepted**. The critical value for the third root is - 26.0, and the hypothesis that the third root is one is **rejected**. Consequently, the S & W procedure accepts the hypothesis that the rank of cointegration is one, and thus, all three procedures yield the same conclusion.

The following observations are prompted by Graphs 3 and 4 below.

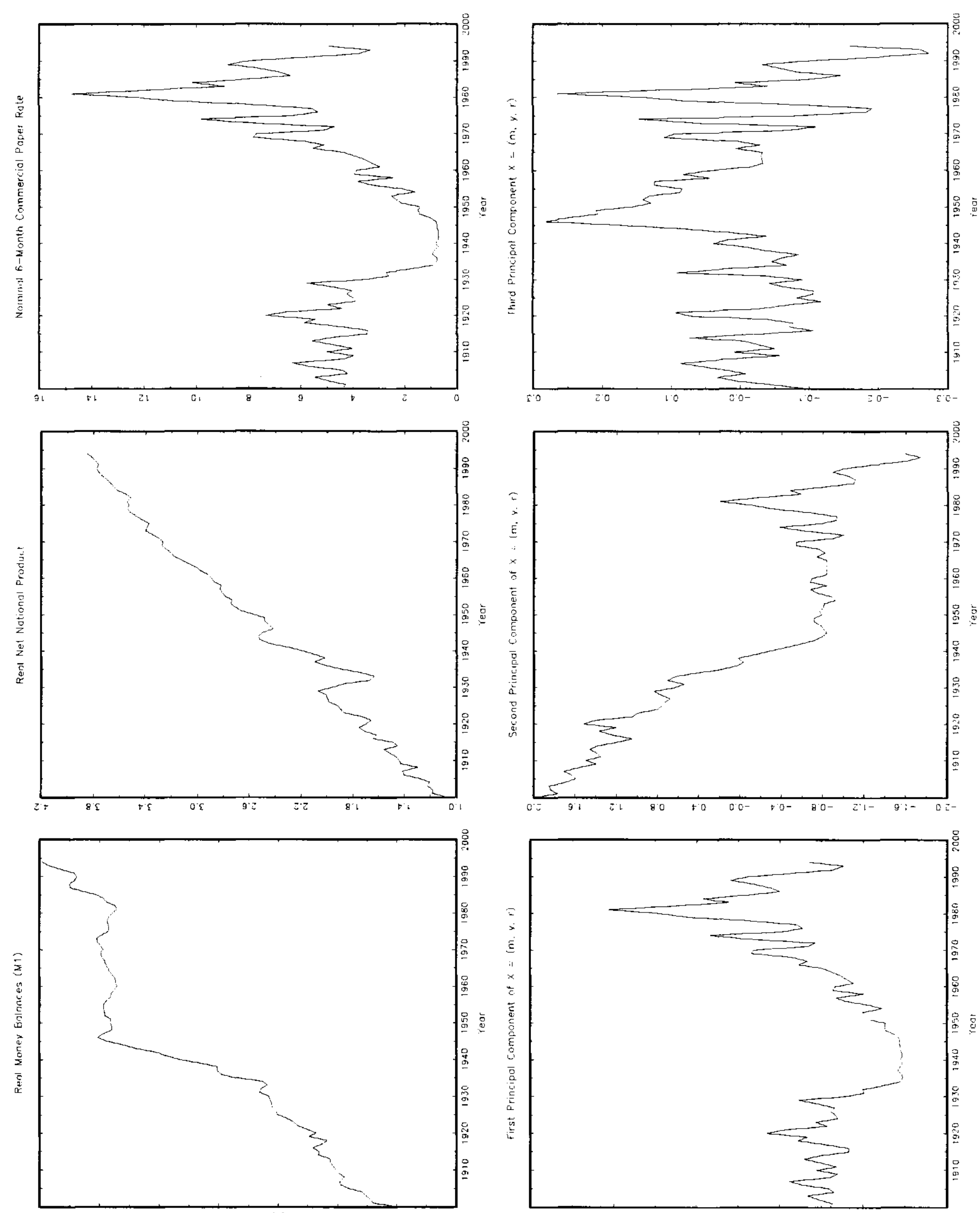
In Graph 3 the upper panel gives the trajectory of the variables of the model. The lower panel gives the trajectory of the three principal components. Notice that the first principal component behaves almost identically with the commercial paper rate. Its range is (-4, 14), while that of the latter is (1, 15). The second component appears to be almost a mirror image of the real demand for money. Its range is (-1.8, 2), while that of the latter is (-0.2, 2.2). The third component, behaves amost like a constant, its range being (-0.275, 0.3). It is quite possible that the finding of cointegration of rank 1, simply relfects the near singularity of the covariance matrix of three stationary processes, and not the intepretation we place on it.

In the upper panel of Graph 4, we exhibit the cointegral vector as defined by the LR procedure; strictly speaking, only the first is a cointegral scalar, the others being irrelevant in the Johansen context. In the lower panel we exhibit the cointegral vector as defined in the conformity context, viz. $\hat{z}_{t\cdot} = X_{t\cdot}\hat{\Pi}(1)$. Notice that the first two elements of $\hat{z_t}.$ have a very compressed range. The third ranges from (-1, 1.8). However, even though its range is extensive, it does not appear to have a "trend". Finally, and in keeping with the nature of $\hat{z}_{t\cdot} = X_{t\cdot}\hat{\Pi}(1)$ as an estimator of a set of linear transformations of the basic cointegral scalar,[13] the shapes of the three trajectories are more similar to each other than those exhibited in the upper panel.

[13] This is so since the rank of cointegration, in the conformity context, has been determined to be one.

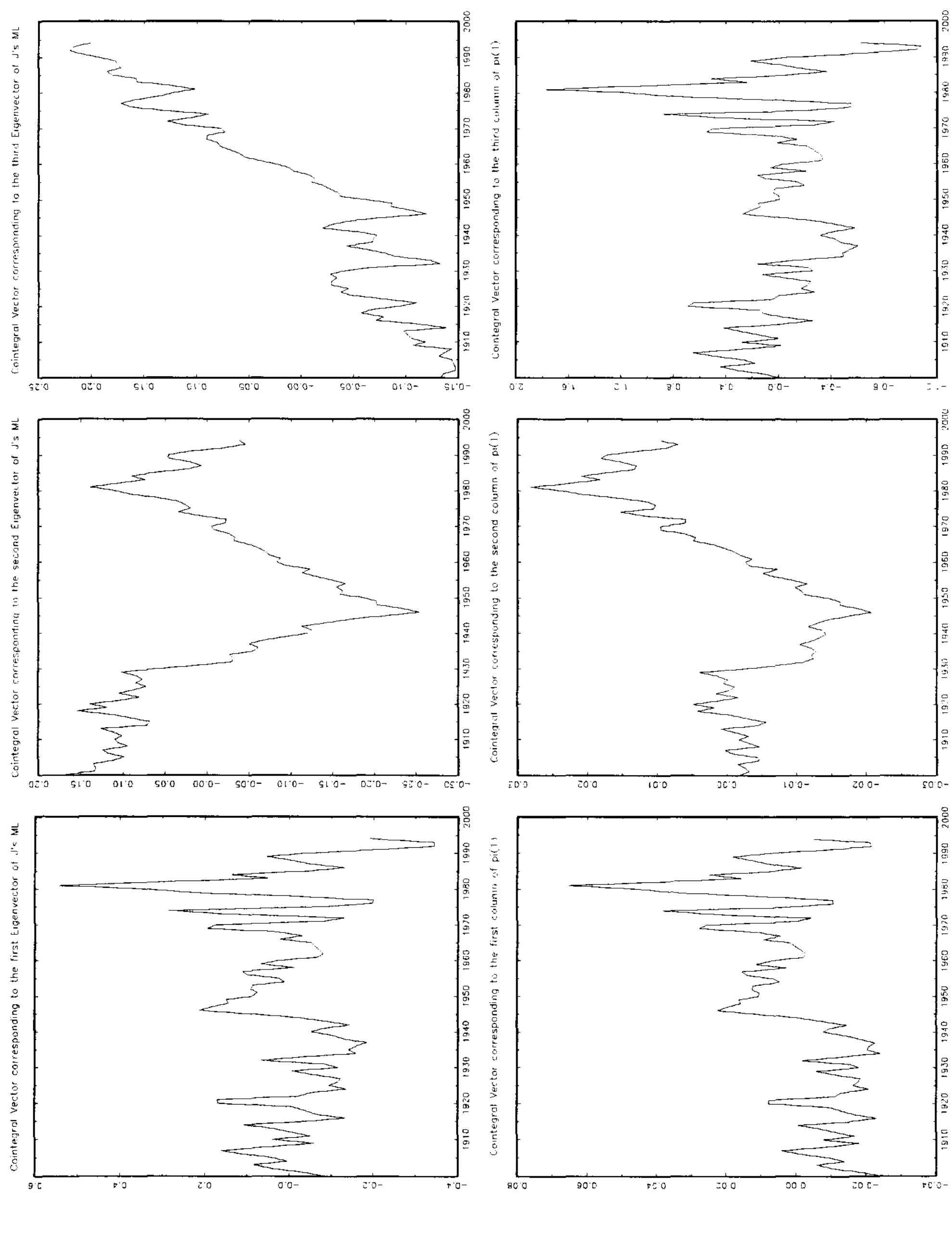
Graph 3

Principal Components and Variables of Money Demand System



Graph 4

Cointegral Vectors defined by the Characteristic Vectors of the LR (Johansen) Procedure and by $X_t.\hat{\Pi}(1)$



REFERENCES

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, New York: Wiley.

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covaraince Matrix Estimation", *Econometrica*, vol. 59, pp. 817-858.

Dhrymes, P. J. (1970), *Econometrics: Statistical Foundations and Applications*, New York: Harper and Row.

Dhrymes, P. J. (1984), *Mathematics for Econometrics*, second edition, New York: Springer Verlag.

Dhrymes, P. J. (1989), *Topics in Advanced Econometrics: vol. I, Probability Foundations*, New York: Springer Verlag.

Dhrymes, P. J. (1995), *Topics in Advanced Econometrics: vol. III, Topics in Time Series*, unpublished.

Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, vol. 74, pp. 42-431.

Dickey, D. A. and W. A. Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, vol.49, pp. 1057-1072.

Engel, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, vol. 55, pp. 251-276.

Epanechnikov, V. A. (1969), "Nonparametric Estimation of a Multivariate Probability Density", *Theory of Probability and its Applications*, vol. 14, pp. 153-158.

Hannan, E. J. (1970), *Multiple Time Series*, New York: Wiley.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, vol. 12, pp. 231-254.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, vol. 59, pp. 1551-1580.

Phillips, P.C.B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica*, vol. 58, pp. 165-93.

Priestley, M. B. (1962), *Spectral Analysis and Time Series*, vols. I and II, New York: Academic Press.

Silverman, B. W. (1986), *Density Esttimation for Statistics and Data Analysis*, London: Chapman and Hall.

Stock, J. H. and M. W. Watson (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, vol. 83, pp. 1097-1107.

Stock, J. H. and M. W. Watson (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica*, vol. 61, pp. 783-820.

Stone, R. A. (1947), "On the Interdependence of Blocks of Transactions", *Journal of the Royal Statistical Society*, Supplement, vol. 9, pp. 1-32.

1995-1996 Discussion Paper Series

Department of Economics
Columbia University
1022 International Affairs Bldg.
420 West 118th Street
New York, N.Y., 10027

The following papers are published in the 1995-96 Columbia University Discussion Paper series which runs from early November to October 31 (Academic Year). Domestic orders for discussion papers are available for purchase at $8.00 (US) each and $140.00 (US) for the series. Foreign orders cost $10.00 (US) for individual paper and $185.00 for the series. To order discussion papers, please send your check or money order payable to Department of Economics, Columbia University to the above address. Be sure to include the series number for the paper when you place an order.

9596-01 Protectionist Response to Import Competition in Declining Industries Reconsidered
J.P. Choi

9596-02 New Estimates on Climate Demand: Evidence from Location Choice
M. Cragg
M. Kahn

9596-03 Enforcement by Hearing
C.W. Sanchirico

9596-04 PreferentialTrading Areas and Multilateralism: Strangers, Friends or Foes?
J. Bhagwati
A. Panagariya

9596-05 Simplification, Progression and a Level Playing Field
William Vickrey

9596-06 The Burden of Proof in Civil Litigation
C. Sanchirico

9596-07 Market Structure and the Timing of Technology Adoption
J.P. Choi
M. Thum

9596-08 The Emergence of the World Economy
R. Findlay

9596-09 The Global Age: From a Skeptical South to a Fearful North
J. Bhagwati

9596-10 A Conformity Test for Cointegration
P.J. Dhrymes

9596-11 Identification and Kullback Information in the GLSEM
P.J. Dhrymes